Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 98 (2007) 317–327
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Estimation for parameters of interest in random effects growth curve models

Wai-Cheung Ip[a,*], Mi-Xia Wu[b], Song-Gui Wang[b], Heung Wong[a]

[a]*Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong*
[b]*College of Applied Sciences, Beijing University of Technology, Beijing 100022, China*

Received 1 March 2005
Available online 8 September 2006

Abstract

In this paper, we consider the general growth curve model with multivariate random effects covariance structure and provide a new simple estimator for the parameters of interest. This estimator is not only convenient for testing the hypothesis on the corresponding parameters, but also has higher efficiency than the least-square estimator and the improved two-stage estimator obtained by Rao under certain conditions. Moreover, we obtain the necessary and sufficient condition for the new estimator to be identical to the best linear unbiased estimator. Examples of its application are given.

AMS 1991 subject classification: 62H12; 62H15; 62F10

Keywords: Growth curve models; Random-effects; Exact test; Best linear unbiased estimator; Maximum likelihood estimator

1. Introduction

The linear mixed model is a popular choice for the treatment of longitudinal data with random effects, see Laird and Ware [6], Diggle et al. [3], Srivastava and VonRosen [14] and Verbeke and Molenberghs [15]. In this paper, we consider the multivariate model in which m distinct characteristics on each of N individuals taken from r different groups are measured on each of p different occasions. The jth characteristic on the ith individual can be assumed to follow the mixed model

$$y_{ij} = X_j \beta_j^{(k)} + Z_j u_{ij} + \varepsilon_{ij}, \quad i = 1, \ldots, N, \quad j = 1, \ldots, m, \quad i \in k\text{th group}, \tag{1}$$

* Corresponding author. Fax: +852 23629045.
E-mail address: mathipwc@polyu.edu.hk (W.C. Ip).

0047-259X/$ - see front matter
doi:10.1016/j.jmva.2006.04.001

where $y_{ij} = (y_{ij1} \cdots y_{ijp})'$, y_{ijl} is the measurement of the jth characteristic at occasion l on the ith individual, X_j and Z_j are, respectively, known $p \times q$ and $p \times c$ design matrices of full column rank, $\beta_j^{(k)}$ is the $q \times 1$ vector of regression parameters on the jth characteristic in the treatment group k $(k = 1, \ldots, r)$, u_{ij} is $c \times 1$ random effect vector, and ε_{ij} is $p \times 1$ random error vector.

If each of the m characteristics that we consider follows a response curve of the same general type over the p occasions, that is, $X_j = X$ and $Z_j = Z$, then for ith individual,

$$y_i = (X \otimes I_m)\beta^{(k)} + (Z \otimes I_m)u_i + \varepsilon_i,$$

where

$$y_i = \mathrm{Vec}((y_{i1}, \ldots, y_{im})'), \quad \beta^{(k)} = \mathrm{Vec}((\beta_1^{(k)}, \ldots, \beta_m^{(k)})'),$$

$$\varepsilon_i = \mathrm{Vec}((\varepsilon_{i1}, \ldots, \varepsilon_{im})'), \quad u_i = \mathrm{Vec}((u_{i1}, \ldots, u_{im})'),$$

$\otimes$ denotes the Kronecker product, and $\mathrm{Vec}(\cdot)$ operator stacks the columns of a matrix below one another to form a column vector. The individual random effects u_i are assumed to be distributed independently as $N(0, \Sigma_u)$, and independent of the error ε_i, with distribution $N(0, I_p \otimes \Sigma_e)$, where Σ_e is $m \times m$ positive definite matrix and Σ_u is $cm \times cm$ nonnegative definite matrix, that is, $\Sigma_e > 0$ and $\Sigma_u \geqslant 0$. Thus the covariance matrix of y_i is

$$\mathrm{Cov}(y_i) = I_p \otimes \Sigma_e + (Z \otimes I_m)\Sigma_u(Z' \otimes I_m) = \Omega \quad \text{(say)}. \tag{2}$$

Let $Y = (y_1, \ldots, y_N)$, $U = (u_1, \ldots, u_N)$, and $E = (\varepsilon_1, \ldots, \varepsilon_N)$, the full multivariate model (1) can be expressed as

$$Y = (X \otimes I_m)\xi A + (Z \otimes I_m)U + E,$$

$$\mathrm{Cov}(\mathrm{Vec}(Y)) = I_N \otimes \Omega, \tag{3}$$

which is a general growth curve model with multivariate random effects covariance structure, where $\xi = (\beta^{(1)}, \ldots, \beta^{(r)})$ is the $qm \times r$ matrix of the growth curve coefficients, and $A = (a_1, \ldots, a_N)$ is the $r \times N$ matrix with full row rank. In particular, if its elements are either 1 or 0 indicating the group from which an observation comes, A is called the ‘group indicator’ matrix in literature. When $m = 1$, model (3) becomes $Y = X\xi A + ZU + E$, and $\Omega = \sigma_e^2 I_p + Z\Sigma_u Z'$, reducing to the single-variable case.

Reinsel [11,12], Azzalini [1], Lange and Laird [7], and Nummi [8,9] considered two special cases: $Z = X$ and $Z = X_c$, where $X = (X_c : X_{\bar{c}})$, and showed that the maximum likelihood estimator (MLE) of ξ is identical to its least-squares estimator (LSE)

$$\hat{\xi} = ((X'X)^{-1}X' \otimes I_m)YA'(AA')^{-1}. \tag{4}$$

However, this is not the case for a general design matrix Z, where the explicit MLE of ξ usually does not exist. Moreover, both the two-stage estimator provided by Khatri [5] and the LSE $\hat{\xi}$ ignore the structure information on the covariance matrix Ω. One alternative is to adopt the improved two-stage estimator suggested by Rao [10] using the structure covariance matrix of the mixed effect model.

In many practical situations, only a part of the parameters of model (3) are meaningful to the researcher. For example, in the rat data of Verbeke and Molenberghs [15], of primary interest is the estimation of changes over time and testing whether these changes are treatment dependent. For the mixed linear model with one random effect, Wu and Wang [17] gave a simple estimator of the part parameter by the reduced model which is often used to remove nuisance parameters. In this paper, we mainly consider this problem under the general growth curve model (3) with random effects.

In Section 2, we introduce a new simple estimator for the part parameter of ξ by a reduced model, which can be superior to the LSE under certain conditions. In Section 3, we consider the optimality of the new estimator, and obtain the necessary and sufficient condition the new estimator to be identical to the best linear unbiased estimator (BLUE). Furthermore, we show that the new estimator is superior to the corresponding two-stage estimator under certain conditions. Examples are presented in Section 4.

2. New estimator

Without loss of generality, we partition $X = (X_1 : X_2)$ in model (3) such that

$$\mathcal{M}(X_1) \subseteq \mathcal{M}(Z), \quad \mathcal{M}(X_2) \cap \mathcal{M}(Z) = \{0\}, \tag{5}$$

where X_1 and X_2 are $p \times l$ and $p \times (q-l)$ matrices, respectively, $(0 \leqslant l \leqslant q)$, and $\mathcal{M}(C)$ is the range space of any matrix C. Partition $\xi = (\xi_1' : \xi_2')'$ conformably. Thus model (3) can be rewritten as

$$Y = (X_1 \otimes I_m)\xi_1 A + (X_2 \otimes I_m)\xi_2 A + (Z \otimes I_m)U + E, \tag{6}$$

where ξ_2 is the parameter matrix of primary interest. In what follows, we mainly consider the estimation of ξ_2.

Denote Q_1 the $p \times (p-c)$ matrix such that $Q_1'Q_1 = I_{p-c}$ and $Q_1'Z = 0$. Thus $Q_1Q_1' = I_p - P_Z$, where $P_Z = Z(Z'Z)^- Z'$ is the orthogonal projector on the space $\mathcal{M}(Z)$. Let $M_Z = I_p - P_Z = Q_1Q_1'$, premultiplying model (6) by $Q_1' \otimes I_m$, we can obtain the reduced model

$$(Q_1' \otimes I_m)Y = (Q_1'X_2 \otimes I_m)\xi_2 A + \varepsilon, \quad \mathrm{Cov}(\mathrm{Vec}(\varepsilon)) = I_{N(p-c)} \otimes \Sigma_e, \tag{7}$$

which is relevant to the matrix parameters ξ_2 and Σ_e. It is readily to see that, under model (7), the MLE of ξ_2 equals to its LSE

$$\tilde{\xi}_2 = \left[(X_2'M_Z X_2)^{-1} X_2' M_Z \otimes I_m\right] Y A'(AA')^{-1}. \tag{8}$$

Let ξ_2 be partitioned as $\xi_2 = (\xi_{21}', \ldots, \xi_{2(q-l)})'$, where ξ_{2i} $(i = 1, \ldots, q-l)$ are $m \times r$ matrices. Assume that the elements of X are functions of the time variable t, such as $X = (f_i(t_j))$, then ξ_{2i} represent the regression coefficients attached to $f_{l+1}(t), \ldots, f_{q-l}(t)$, for the m characteristics and r groups. Denote $\Psi = (\xi_{21}, \ldots, \xi_{2(q-l)})$. Applying definition of the restricted maximum likelihood (REML) estimator, we can obtain the REML estimator of Σ_e under reduced model (7), that is,

$$\tilde{\Sigma}_e = \sum_{i=1}^{N} \left(Y_i Q_1 - \tilde{\Psi}(X_2'Q_1 \otimes a_i)\right)\left(Y_i Q_1 - \tilde{\Psi}(X_2'Q_1 \otimes a_i)\right)' \Big/ k, \tag{9}$$

where $k = N(p-c) - (q-l)r$, $Y_i = (y_{i1}, \ldots, y_{im})'$, and $\tilde{\Psi} = (\tilde{\xi}_{21}, \ldots, \tilde{\xi}_{2(q-l)})$.

Under the original model (6), estimators $\tilde{\xi}_2$ and $\tilde{\Sigma}_e$ are also unbiased for ξ_2 and Σ_e, respectively.

Theorem 2.1.
(a) $\tilde{\xi}_2 \sim N(\xi_2, (AA')^{-1} \otimes (X_2' M_Z X_2)^{-1} \otimes \Sigma_e)$,
(b) $k \cdot \tilde{\Sigma}_e \sim W_m(k, \Sigma_e)$, *and independent of* $\tilde{\xi}_2$,

where $W_m(k, \Sigma_e)$ *is the* m*-dimensional Wishart distribution with* k *degrees of freedom and parameter matrix* Σ_e.

Proof. (a) is obvious so we only need to prove (b). Noting $\mathrm{Vec}(AXC) = (C' \otimes A)\,\mathrm{Vec}(X)$, thus model (7) can be rewritten as

$$Y_i Q_1 = \Psi(X_2' Q_1 \otimes a_i) + E_i, \quad \mathrm{Cov}(\mathrm{Vec}(E_i)) = I_{p-c} \otimes \Sigma_e, \quad i = 1, \ldots, N, \tag{10}$$

where Y_i is the same as that in (9), $\Psi = (\xi_{21}, \ldots, \xi_{2(q-1)})$, and $E_i = (\varepsilon_{i1}, \ldots, \varepsilon_{im})' Q_1$.
Denote

$$Y_0=(Y_1 Q_1, \ldots, Y_N Q_1), \quad X_0=(X_2' Q_1 \otimes a_1, \ldots, X_2' Q_1 \otimes a_N), \quad E_0=(E_1, \ldots, E_N),$$

then model (7) can also be rewritten as

$$Y_0 = \Psi X_0 + E_0, \quad \mathrm{Cov}(\mathrm{Vec}(E_0)) = I_{N(p-c)} \otimes \Sigma_e. \tag{11}$$

Thus we get the equivalent forms of $\tilde{\Psi} = (\tilde{\xi}_{21}, \ldots, \tilde{\xi}_{2(q-l)})$ and $\tilde{\Sigma}_e$, respectively,

$$\begin{aligned}
\tilde{\Psi} &= Y_0 X_0' (X_0 X_0')^{-1} = \sum_{i=1}^{N} Y_i Q_1 \left(Q_1' X_2 (X_2' M_Z X_2)^{-1} \otimes a_i' (AA')^{-1} \right), \\
\tilde{\Sigma}_e &= Y_0 (I_{N(p-c)} - X_0' (X_0 X_0')^{-1} X_0) Y_0' / k \\
&= \left(\sum_{i=1}^{N} Y_i M_Z Y_i' - \tilde{\Psi} (X_2' M_Z X_2 \otimes AA') \tilde{\Psi}' \right) \Big/ k.
\end{aligned}$$

It follows readily from (11) that $\tilde{\Sigma}_e$ is independent of $\tilde{\Psi}$ and $k \cdot \tilde{\Sigma}_e \sim W_m(k, \Sigma_e)$. The proof of Theorem 2.1 is completed. □

Remark 2.1. $\mathrm{Cov}(\mathrm{Vec}(\tilde{\xi}_2))$ does not depend on the matrix Σ_u.

According to Theorem 2.1, the new estimator $\tilde{\xi}_2$ can be used to construct an exact test on Ψ, i.e. on ξ_2 for the general linear hypothesis $\mathrm{H}_0 : L\Psi G = 0$. In fact, the Wilks's Λ is given by

$$\Lambda = \frac{|W|}{|W + H|}, \tag{12}$$

where $W = k(L\tilde{\Sigma}_e L')$ and $H = (L\tilde{\Psi}G) \left[G'(X_2' M_Z X_2 \otimes AA')^{-1} G \right]^{-1} (G' \tilde{\Psi}' L')$.
On the other hand, from (4), it is easy to obtain that the LSE of ξ_2 under model (6) is

$$\hat{\xi}_2 = \left((X_2' M_{X_1} X_2)^{-1} X_2' M_{X_1} \otimes I_m \right) Y A' (AA')^{-1},$$

and

$$\begin{aligned}
\mathrm{Cov}(\mathrm{Vec}(\hat{\xi}_2)) = (AA')^{-1} \otimes \Big[&(X_2' M_{X_1} X_2)^{-1} \otimes \Sigma_e + \left((X_2' M_{X_1} X_2)^{-1} X_2' M_{X_1} Z \otimes I_m \right) \Sigma_u \\
&\times \left(Z' M_{X_1} X_2 (X_2' M_{X_1} X_2)^{-1} \otimes I_m \right) \Big].
\end{aligned} \tag{13}$$

Combining with Remark 2.1, it is reasonable to expect that $\tilde{\xi}_2$ may be superior to $\hat{\xi}_2$ under some conditions. Suppose that X_1, X_2 and Z in (5) satisfy the following conditions:

$$\begin{aligned} &X_1'X_2 = 0, \quad Z = (X_1, Z_0), \\ &\mathcal{M}(Z_0) \cap \mathcal{M}(X_i) = \{0\}, \quad i = 1, 2, \end{aligned} \tag{14}$$

then $M_Z X_1 = 0$, $M_{X_1} X_2 = X_2$ and

$$M_Z = M_{X_1} - M_{X_1} Z_0 (Z_0' M_{X_1} Z_0)^{-1} Z_0' M_{X_1}. \tag{15}$$

For the proof of the last equality see [13]. By the use of (14) and the matrix identity

$$(A + BCB')^{-1} = A^{-1} - A^{-1}B(B'A^{-1}B + C^{-1})^{-1}B'A^{-1},$$

we can obtain that

$$(X_2' M_Z X_2)^{-1} = (X_2'X_2)^{-1} + (X_2'X_2)^{-1} X_2' Z_0 (Z_0' M_X Z_0)^{-1} Z_0' X_2 (X_2'X_2)^{-1}.$$

Thus $\mathrm{Cov}(\mathrm{Vec}(\tilde{\xi}_2))$ can be rewritten as

$$\begin{aligned} \mathrm{Cov}(\mathrm{Vec}(\tilde{\xi}_2)) = (AA')^{-1} \otimes \Big[&(X_2'X_2)^{-1} \otimes \Sigma_e + \left((X_2'X_2)^{-1} X_2' Z_0 \otimes I_m \right) \\ &\times \left((Z_0' M_X Z_0)^{-1} \otimes \Sigma_e \right) \left(Z_0' X_2 (X_2'X_2)^{-1} \otimes I_m \right) \Big]. \end{aligned} \tag{16}$$

Furthermore, under the assumption (14), (13) can be simplified as

$$\begin{aligned} \mathrm{Cov}(\mathrm{Vec}(\hat{\xi}_2)) = (AA')^{-1} \otimes \Big[&(X_2'X_2)^{-1} \otimes \Sigma_e \\ &+ \left((X_2'X_2)^{-1} X_2' Z_0 \otimes I_m \right) (\Sigma_u)_{22} \left(Z_0' X_2 (X_2'X_2)^{-1} \otimes I_m \right) \Big], \end{aligned} \tag{17}$$

where $(\Sigma_u)_{22} = (0, I_{m(c-l)}) \Sigma_u (0, I_{m(c-l)})'$ is $m(c-l) \times m(c-l)$ nonnegative definite matrix.

Comparing (16) with (17), we obtain the condition for which the new estimator $\tilde{\xi}_2$ is superior to the LSE $\hat{\xi}_2$ in the following theorem.

Theorem 2.2. *If conditions* (14) *hold, then* $\mathrm{Cov}(\mathrm{Vec}(\tilde{\xi}_2)) < \mathrm{Cov}(\mathrm{Vec}(\hat{\xi}_2))$ *if and only if*

$$X_2' Z_0 \neq 0 \quad \textit{and} \quad ((Z_0' M_X Z_0)^{-1} \otimes \Sigma_e) < (\Sigma_u)_{22}. \tag{18}$$

For example, $X_1 = (1, 1, 1)'$, $X_2 = (-1, 0, 1)'$, $Z_0 = (1, 4, 9)'$, *and* $(\Sigma_u)_{22} = a \cdot \Sigma_e$, *then* $\tilde{\xi}_2$ *is superior to* $\hat{\xi}_2$ *only if* $a > \frac{3}{2}$.

Remark 2.2. The condition (18) is irrelevant to the other sub-matrices of Σ_u besides $(\Sigma_u)_{22}$.

In the following section, we will show that the new estimator $\tilde{\xi}_2$ can also be superior to the improved two-stage estimator of ξ_2 under some conditions.

3. The optimality of the new estimator

In this section, we consider the optimality of the new estimator $\tilde{\xi}_2$ of ξ_2, the parameter matrix of primary interest, and give the necessary and sufficient condition under which $\tilde{\xi}_2$ is equal to the BLUE of ξ_2 for the original model (6).

Theorem 3.1. *$\tilde{\xi}_2$ is the BLUE of ξ_2 under model* (6), *if and only if*

$$Z'M_{X_1}X_2 = 0, \tag{19}$$

where $M_{X_1} = I_p - X_1(X_1'X_1)^{-1}X_1'$.

Proof. By Lemma 5.4.3 [16], $\tilde{\xi}_2$ is the BLUE of ξ_2 under model (6) if and only if

$$\mathrm{Cov}(\mathrm{Vec}(\tilde{\xi}_2), (I_n - P_{A'} \otimes P_X \otimes I_m)\,\mathrm{Vec}(Y)) = 0, \tag{20}$$

where $n = Npm$, that is,

$$(AA')^{-1}A \otimes ((X_2'M_Z X_2)^{-1}X_2'M_Z \otimes I_m)\Omega(M_X \otimes I_m) = 0,$$

which is equivalent to

$$X_2'M_Z M_X = 0. \tag{21}$$

From (5), we have $X_1'M_Z = 0$, thus

$$(21) \Leftrightarrow X'P_Z M_X = 0 \Leftrightarrow P_X P_Z M_X = 0 \Leftrightarrow P_X P_Z = P_{X_1}.$$

Combining Theorem 1 (A10, A11) of [2] and the fact that

$$P_X = P_{X_1} + M_{X_1}X_2(X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1}. \tag{22}$$

Eq. (21) can been simplified to $Z'M_{X_1}X_2 = 0$. The proof of the Theorem 3.1 is completed. □

The BLUE of ξ_2 under model (6) is

$$\xi_2^* = C_{22.1}^{-1}C_2 M_{C_1'}\Omega^{-1/2}YA'(AA')^{-1}, \tag{23}$$

where $C_i = (X_i' \otimes I_m))\Omega^{-1/2}, i = 1, 2$, $C_{22.1} = C_2 M_{C_1'} C_2'$, $M_{C_1'} = I - C_1'(C_1'C_1)^{-1}C_1$. Obviously, ξ_2^* is an usually unfeasible estimator since Ω in (23) includes two unknown covariance matrices Σ_u and Σ_e. However, if condition (19) holds, then by Theorem 3.1, we have $\xi_2^* = \tilde{\xi}_2$, that means the existence of the explicit ML estimator of the part parameter ξ_2.

Theorem 3.2. *Under model* (6), *the following statements are equivalent*:

(a) $Z'M_{X_1}X_2 = 0$,
(b) $\tilde{\xi}_2 = \hat{\xi}_2$,
(c) $\hat{\xi}_2 = \xi_2^*$.

Proof. (i) Proof of $(a) \Leftrightarrow (c)$. Similar to the proof of Theorem 3.1, we have

$$\hat{\xi}_2 = \xi_2^* \Leftrightarrow \mathrm{Cov}(\mathrm{Vec}(\hat{\xi}_2), (I_n - P_{A'} \otimes P_X \otimes I_m)\,\mathrm{Vec}(Y)) = 0,$$

which can be simplified as

$$(X_2'M_{X_1}Z \otimes I_m)\Sigma_u(Z'M_X \otimes I_m) = 0. \tag{24}$$

Since $\Sigma_u \geqslant 0$ is arbitrary, (24) is equivalent to $Z'M_{X_1}X_2 = 0$ or $Z'M_X = 0$. From (5), we have

$$Z'M_X = 0 \Leftrightarrow \mathcal{M}(Z) = \mathcal{M}(X_1) \Rightarrow Z'M_{X_1}X_2 = 0,$$

thus (24) is equivalent to $Z'M_{X_1}X_2 = 0$, that is, (a) $\Leftrightarrow$ (c).

(ii) Proof of (a) $\Leftrightarrow$ (b). Obviously, (a) $\Rightarrow$ (b). Now, we consider (b) $\Rightarrow$ (a). If (b) holds, then

$$(X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1} = (X_2'M_ZX_2)^{-1}X_2'M_Z. \tag{25}$$

Postmultiplying (25) by Z, we have $(X_2'M_{X_1}X_2)^{-1}X_2'M_{X_1}Z=0$, which is equivalent to $Z'M_{X_1}X_2 = 0$. The proof of Theorem 3.2 is completed. $\square$

Theorem 3.2 shows that the new estimator $\tilde{\xi}_2$ and the LSE $\hat{\xi}_2$ can achieve optimality simultaneously, and both necessary and sufficient conditions are $Z'M_{X_1}X_2 = 0$.

Remark 3.1. The condition $Z'M_{X_1}X_2 = 0$ is weaker than the necessary and sufficient condition for $\hat{\xi} = \xi^*$ under model (6).

In fact, by Lemma 5.4.3 of [16], we can testify that $\hat{\xi} = \xi^*$ if and only if $Z'X = 0$ or $\mathcal{M}(Z) = \mathcal{M}(X_1)$, which are two special cases for $Z'M_{X_1}X_2 = 0$.

Remark 3.1 sharpens the intuition that while the explicit MLE of the whole parameter matrix does not exist, it may do so for the part parameter matrix.

For the covariance matrix with random effects such as (2), Rao [10] presented an improved two-stage estimator of ξ

$$\hat{\xi}_{T_1} = \hat{\xi} - ((X'X)^{-1}X' \otimes I_m)SQ(Q'SQ)^{-1}Q'YA'(AA')^{-1},$$

where $Q = M_XZ \otimes I_m$, $T_1 = Q'Y$, and $S = Y(I_n - P_{A'})Y'$. $\hat{\xi}_{T_1}$ is derived by covariance adjustment in the LSE $\hat{\xi}$ using the concomitant variable T_1, Rao [10] and Grizzle and Allen [4] proved

$$\begin{aligned}\mathrm{Cov}(\mathrm{Vec}(\hat{\xi}_{T_1})) &= \frac{N-r-1}{N-r-mk_0-1}(AA')^{-1} \otimes \{(X' \otimes I_m)\Omega^{-1}(X \otimes I_m)\}^{-1} \\ &= \frac{N-r-1}{N-r-mk_0-1}\mathrm{Cov}(\mathrm{Vec}(\xi^*)) \geqslant \mathrm{Cov}(\mathrm{Vec}(\xi^*)),\end{aligned} \tag{26}$$

where $k_0 = \mathrm{rk}(M_XZ)$. Clearly, (26) takes the equality if and only if $k_0 = 0$, which requires $\mathcal{M}(Z) \subseteq \mathcal{M}(X)$. In this case the improved two-stage estimator $\hat{\xi}_{T_1}$ is equal to the LSE $\hat{\xi}$, and $\tilde{\xi}_2 = \hat{\xi}_2 = \hat{\xi}_{2T_1}$, where $\hat{\xi}_{2T_1}'$ is the corresponding improved two-stage estimator of ξ_2, $\hat{\xi}_{T_1} = (\hat{\xi}_{1T_1}', \hat{\xi}_{2T_1}')'$. In the following corollary, we will give a set of conditions for the new estimator $\tilde{\xi}_2$ to be superior to the improved two-stage estimator $\hat{\xi}_{2T_1}$ under the case $k_0 \neq 0$.

Corollary 3.1. *If $\mathcal{M}(X) \cap \mathcal{M}(Z) = \mathcal{M}(X_1)$, $\mathcal{M}(Z) \neq \mathcal{M}(X_1)$ and $Z'M_{X_1}X_2 = 0$, then*

$$\mathrm{Cov}(\mathrm{Vec}(\hat{\xi}_{2T_1})) > \mathrm{Cov}(\mathrm{Vec}(\xi_2^*)) = \mathrm{Cov}(\mathrm{Vec}(\tilde{\xi}_2)) = \mathrm{Cov}(\mathrm{Vec}(\hat{\xi}_2)). \tag{27}$$

The conditions $\mathcal{M}(X) \cap \mathcal{M}(Z) = \mathcal{M}(X_1)$ and $\mathcal{M}(Z) \neq \mathcal{M}(X_1)$ ensure $k_0 \neq 0$. Upon use of Theorem 3.1 and (26), we can obtain (27).

In order to understand the set of conditions in Corollary 3.1, without loss of generality, we assume $Z = (X_1 : Z_0)$. By (5), we have

$$\mathcal{M}(Z_0) \cap \mathcal{M}(X) = \{0\},$$

and

$$Z'M_{X_1}X_2 = 0 \quad \Leftrightarrow \quad Z_0'M_{X_1}X_2 = 0.$$

According to Theorem 1 (A10, A11) of [2], $Z_0' M_{X_1} X_2 = 0$ is equivalent to $P_{Z_0} P_X = 0$, thus

$$Z' M_{X_1} X_2 = 0 \quad \Leftrightarrow \quad Z_0' X = 0, \tag{28}$$

and hence the conditions in Corollary 3.1 are equivalent to $Z_0' X = 0$, and $Z_0 \neq 0$. For example, $X_1 = (1, 1, 1, 1)'$, $X_2 = (1, 2, 3, 4)'$, $Z = (X_1, Z_0)$, where $Z_0 = (1, -1, -1, 1)'$, clearly, $Z_0 \neq 0$ and $Z_0' X = 0$.

4. Examples

In this section, we shall give two simple examples to illustrate the foregoing results.

Example 1. Consider the model for the rat data (see Verbeke and Molenberghs [14]):

$$y_{ij} = \begin{cases} \beta_{0_1} + u_{0j} + \beta_1 t_i + u_{1j} z_i + \varepsilon_{ij} & \text{if low dose,} \\ \beta_{0_2} + u_{0j} + \beta_2 t_i + u_{1j} z_i + \varepsilon_{ij} & \text{if high dose,} \\ \beta_{0_3} + u_{0j} + \beta_3 t_i + u_{1j} z_i + \varepsilon_{ij} & \text{if control,} \end{cases} \tag{29}$$

where y_{ij} is the observation for the jth individual at t_i time point, $i = 1, \ldots, p$, $u_j = (u_{0j}, u_{1j})'$ is random effect vector with normal distribution $N(0, \Sigma_u)$, and random error ε_{ij} follows the distribution $N(0, \sigma_e^2)$. We assume that $u_1, \ldots, u_N, \varepsilon_{11}, \ldots, \varepsilon_{1N}, \ldots, \varepsilon_{p1}, \ldots, \varepsilon_{pN}$ are independent each other. Of primary interest is the estimation of the slopes β_1, β_2 and β_3, and testing whether these slopes are equal to each other.

Without loss of generality, we assume that $Y_1 = (y_1, \ldots, y_{n_1})$, $Y_2 = (y_{n_1+1}, \ldots, y_{n_2}$ and $Y_3 = (y_{n_2+1}, \ldots, y_N)$ are the matrices of the observations for the low-dose group, high-dose group and control group, respectively, where $y_j = (y_{1j}, \ldots, y_{pj})'$. Denote

$$Y = (Y_1 : Y_2 : Y_3), \quad U = (u_1, \ldots, u_N), \quad T = (t_1, \ldots, t_p)',$$

$$Z_0 = (z_1, \ldots, z_p)', \quad X = (\mathbf{1}_p : T), \quad Z = (\mathbf{1}_p : Z_0)$$

and

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \beta_{01} & \beta_{02} & \beta_{03} \\ \beta_1 & \beta_2 & \beta_3, \end{pmatrix}, \quad A = \begin{pmatrix} \mathbf{1}_{n_1}' & 0_{1\times(n_2-n_1)} & 0_{1\times(N-n_2)} \\ 0_{1\times n_1} & \mathbf{1}_{(n_2-n_1)}' & 0_{1\times(N-n_2)} \\ 0_{1\times n_1} & 0_{1\times(n_2-n_1)} & \mathbf{1}_{(N-n_2)}' \end{pmatrix},$$

then model (4.2) can be rewritten as

$$Y = X\xi A + ZU + E = \mathbf{1}_p \xi_1 A + T\xi_2 A + ZU + E. \tag{30}$$

The covariance matrix of Y is

$$\mathrm{Cov}(\mathrm{Vec}(Y)) = I_N \otimes (Z\Sigma_u Z' + \sigma_e^2 I_p).$$

Denote Q_1 the $p \times (p-2)$ matrix, such that $Q_1 Q_1' = M_Z = I_p - P_Z$, then the reduced model of (30) may be represented by

$$Q_1' Y = Q_1' T \xi_2 A + Q_1' E, \quad \mathrm{Cov}(\mathrm{Vec}(Q_1' Y)) = I_N \otimes (\sigma_e^2 I_{p-2}). \tag{31}$$

By (8) and (9), we obtain the new estimators of $\xi_2 = (\beta_1, \beta_2, \beta_3)$ and σ^2

$$\tilde{\xi}_2 = \frac{1}{T' M_Z T} T' M_Z (\overline{Y}_1 : \overline{Y}_2 : \overline{Y}_3),$$

or

$$\tilde{\beta}_r = \frac{1}{T'M_Z T} T'M_Z \overline{Y}_r, \quad r = 1, 2, 3, \tag{32}$$

and

$$\tilde{\sigma}_e^2 = \sum_{r=1}^{3} SS_r/k, \tag{33}$$

where $k = N(p-2) - 3$, $\overline{Y}_r = (\overline{Y}_{1r}, \ldots, \overline{Y}_{pr})'$,

$$\overline{Y}_{i1} = \sum_{j=1}^{n_1} y_{ij}/n_1, \quad \overline{Y}_{i2} = \sum_{j=n_1+1}^{n_2} y_{ij}/(n_2 - n_1), \quad \overline{Y}_{i3} = \sum_{j=n_2+1}^{N} y_{ij}/(N - n_2),$$

$$SS_1 = \sum_{j=1}^{n_1} (y_j - \tilde{\beta}_1 T)' M_Z (y_j - \tilde{\beta}_1 T),$$

$$SS_2 = \sum_{j=n_1+1}^{n_2} (y_j - \tilde{\beta}_2 T)' M_Z (y_j - \tilde{\beta}_2 T), \quad \text{and}$$

$$SS_3 = \sum_{j=n_2+1}^{N} (y_j - \tilde{\beta}_3 T)' M_Z (y_j - \tilde{\beta}_3 T).$$

By Theorem 2.1, $\tilde{\xi}_2$ and $\tilde{\sigma}^2$ are independent, based on which we can construct a test statistic for testing $\mathrm{H}_0 : \beta_1 = \beta_2 = \beta_3$. Let

$$H' = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix},$$

then H_0 is equivalent to testing $\xi_2 H = 0$. The exact test statistic is

$$F = \frac{\tilde{\xi}_2 H (H'(AA')^{-1} H)^{-1} H' \tilde{\xi}_2'/2}{\tilde{\sigma}_e^2} (T' M_Z T),$$

which is an F-statistic and hence $F \sim F_{2,k}$ if $\xi_2 H = 0$ holds.

By Theorem 2.2, under the case $T'Z_0 \neq 0$, if the variance of the error σ^2 satisfies

$$\sigma_e^2 < (Z_0' M_X Z_0)(\Sigma_u)_{22} = (Z_0' M_X Z_0)\,\mathrm{Var}(u_{1j}), \tag{34}$$

then $\tilde{\xi}_2$ is superior to the LSE of ξ_2

$$\hat{\xi}_2 = \frac{1}{(T - \bar{t} \cdot \mathbf{1}_p)'(T - \bar{t} \cdot \mathbf{1}_p)} (T - \bar{t} \cdot \mathbf{1}_p)' (\overline{Y}_1 : \overline{Y}_2 : \overline{Y}_3). \tag{35}$$

By Theorem 3.2, if $T'Z_0 = 0$ and $\mathbf{1}_p Z_0 = 0$, then $\tilde{\xi}_2 = \hat{\xi}_2$, and $\tilde{\xi}_2$ is the BLUE of ξ_2 under model (29). Furthermore, if $Z_0 \neq 0$, by Corollary 3.1, then $\tilde{\xi}_2$ is superior to the improved two-stage estimator of ξ_2.

Example 2. Consider the model for systolic and diastolic blood pressure data:

$$y_{ijl} = \begin{cases} \beta_{0_j}^{(1)} + u_{ij} + \beta_j^{(1)} x_l + v_{ij} t_l + \varepsilon_{ijl} & \text{if treatment 1 group,} \\ \beta_{0_j}^{(2)} + u_{ij} + \beta_j^{(2)} x_l + v_{ij} t_l + \varepsilon_{ijl} & \text{if treatment 2 group,} \end{cases} \tag{36}$$

$$i = 1, \ldots, N, \quad j = 1, 2, \quad l = 1, \ldots, p,$$

where y_{i1l}, y_{i2l} are the systolic and diastolic blood pressure for the ith patients (with moderate essential hypertension) at t_l time point, respectively, $\beta_{0_j}^{(1)}$ and $\beta_{0_j}^{(2)}$ are fixed intercepts, $\beta_j^{(1)}$ and $\beta_j^{(2)}$ are the fixed effects of treatment 1 and 2 on the endpoints, respectively. We assume that random individual effect $u_i = (u_{i1}, u_{i2}, v_{i1}, v_{i2})'$ follows distribution $N(0, \Sigma_u)$, and independent of the error $\varepsilon_i = (\varepsilon_{i11}, \varepsilon_{i21}, \ldots, \varepsilon_{i1p}, \varepsilon_{i2p})'$ with distribution $N(0, I_p \otimes \Sigma_e)$.

Clearly, (36) is the case of model (3) with $m = 2$,

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \beta_{01}^{(1)} & \beta_{01}^{(2)} \\ \beta_{02}^{(1)} & \beta_{02}^{(2)} \\ \hline \beta_1^{(1)} & \beta_1^{(2)} \\ \beta_2^{(1)} & \beta_2^{(2)} \end{pmatrix}, \quad A = \begin{pmatrix} \mathbf{1}_{n_1}' & 0_{1\times(N-n_1)} \\ 0_{1\times n_1} & \mathbf{1}_{(N-n_1)}' \end{pmatrix}, \tag{37}$$

and

$$X = (\mathbf{1}_p, x), \quad Z = (\mathbf{1}_p, t), \quad U = (u_1, \ldots, u_N).$$

Here $x = (x_1, \ldots, x_p)'$ and $t = (t_1, \ldots, t_p)$ are designs vectors on dose and time point.

Of primary interest is the testing on ξ_2, the effects of treatment 1 and 2 on systolic and diastolic blood pressure. According to Theorem 2.1, we can obtain the actual test statistic on ξ_2, Wilks's Λ statistic (12), which is constructed by the new estimator $\tilde{\xi}_2$ given by (8).

Otherwise, the new estimator of ξ_2 is superior to the LSE and two-stage estimator under the conditions of Theorem 2.2 and Corollary 3.1. See the following typical designs on dose x and time point t.

Take $x = (1, 1, 0, -1, -1)$, $t = (1, 2, 3, 4, 5)$, then conditions (18) can be equivalent to

$$\Sigma_e = \mathrm{Cov}((\varepsilon_{i1l}, \varepsilon_{i2l})') < \mathrm{Cov}((v_{i1}, v_{i2})'). \tag{38}$$

That is, if the covariance matrix of error is lower than the covariance matrix of time effect under Löwner partial ordering, then the new estimator of ξ_2 is superior to the LSE.

Take $x = (0, 2, 2, 2, 0)$, $t = (-2, -1, 0, 1, 2)$, it is easy to verify that

$$Z' M_{X_1} x = (0, t'x)' = (0, 0)',$$

so the new estimator of ξ_2 is superior to the two-stage estimator provided by Rao.

The above two examples show that Theorem 2.2 and Corollary 3.1 are helpful for the study on optimal design of experiments too.

Acknowledgments

This research was supported by a grant from The Hong Kong Polytechnic University Research Committee. Wu's and Wang's research was also partially supported by National Natural Science

Foundation of China. We are grateful to the two referees for their detailed suggestions which considerably improved the quality of the paper.

References

[1] A. Azzalini, Growth curves analysis for patterned covariance matrices, in: M.L. Puri, J.P. Vilaplana, W. Wertz (Eds.), New Perspectives in Theoretical and Applied Statistics, Wiley, New York, 1987, pp. 61–74.
[2] J.K. Baksalary, Algebraic characterizations and statistical implications of the commutativity of orthogonal projectors, in: Tarmo Pukkila, S. Puntanen (Eds.), Proceedings of the Second International Tampere Conference in Statistics, 1987, pp. 113–142.
[3] P.J. Diggle, K.E. Liang, S.L. Zeger, Analysis of Longitudinal Data, Oxford Science, New York, 1994.
[4] J.E. Grizzle, D.M. Allen, Analysis of growth and dose response curves, Biometrics 25 (1969) 357–381.
[5] C.C. Khatri, A note on a MANOVA model applied to problems in growth curves, Ann. Inst. Statist. Math. 18 (1966) 75–86.
[6] N.M. Laird, J.H. Ware, Random effects models for longitudinal data, Biometrics 38 (1982) 963–974.
[7] N. Lange, N.M. Laird, The effect of covariance structure on variance estimation in balanced growth-curve models with random parameters, J. Amer. Statist. Assoc. 84 (1989) 241–247.
[8] T. Nummi, Estimation in a random effects growth curve model, J. Appl. Statist. 24 (1997) 157–168.
[9] T. Nummi, Möttönen, On the analysis of multivariate growth curves, Metrika 52 (2000) 77–89.
[10] C.R. Rao, Least squares theory using an estimated dispersion matrix and its application to measurement of signals, in: L. Lecam, J. Neyman (Eds.), Proceeding of the Fifth Berkeley Symposium on Math Statistics & Probability, first ed., 1967, pp. 355–372.
[11] G. Reinsel, Multivariate repeated-measurement or growth curve models with multivariate random-effects covariance structure, J. Amer. Statist. Assoc. 77 (1982) 190–195.
[12] G. Reinsel, Estimation and prediction in a multivariate random effects generalized linear model, J. Amer. Statist. Assoc. 79 (1984) 406–414.
[13] G.A.F. Seber, Linear Regression Analysis, Wiley, New York, 1977 p. 66.
[14] M.S. Srivastava, D. VonRosen, Growth curve models, in: S. Ghosh (Ed.), Multivariate Analysis, Design of Experiments, and Survey Sampling, Marcel Dekker Inc., New York, 1999.
[15] G. Verbeke, G. Molenberghs, Linear Mixed Models for Longitudinal Data, Springer, New York, 2000.
[16] S.G. Wang, S.C. Chow, Advanced Linear Models, Marcel Dekker Inc., New York, 1994.
[17] M.X. Wu, S.G. Wang, Parameter estimation in a partitioned mixed-effects model, Chinese J. Eng. Math. 19 (2) (2002) 31–38.